To appear in Proceedings of the RoboCup International Conference, 2021.

Towards building rapport with a Human Support Robot

Katarzyna Pasternak[1], Zishi Wu[1], Ubbo Visser[1], and Christine Lisetti[2]

[1] University of Miami, Coral Gables, FL 33146
Department of Computer Science
{kwp|zishi|visser}@cs.miami.edu
[2] Florida International University, 11200 SW 8th St, Miami, FL 33199
Knights Foundation School of Computing and Information Sciences
lisetti@cis.fiu.edu

Abstract. Human support robots (mobile robots able to perform useful domestic manipulative tasks) might be better accepted by people if they can communicate in ways they naturally understand: e.g. speech, but also facial expressions, postures, among others. Subtle (unconscious) mirroring of nonverbal cues during conversations promotes rapport building, essential for good communication. We investigate whether, as in human-human communication, the ability of a robot to mirror its user's head movements and facial expressions in real time can improve the user's experience with it. We describe the technical integration of a Toyota Human Support Robot (HSR) with a facially expressive 3D embodied conversational agent (ECA) (named ECA-HSR). The HSR and the ECA are aware of the user's head movements and facial emotions, and can mirror them, in real time. We then discuss a user study we designed in which participants interacted with ECA-HSR in a simple social dialog task with three conditions: mirroring of user's head movements, mirroring of user's facial emotions, and mirroring of both user's head movements and facial emotions. Our results suggest that interacting with an ECA-HSR that mirrors both the user's head movements and the facial expressions is preferred over the other conditions. Among other insights, the study revealed that the accuracy of open source, real-time recognition of facial expressions of emotion needs improvement for the best user's acceptance.

Keywords: Human Robot Interaction · 3D Embodied Conversational Agents · Autonomous Support Robots · Nonverbal Communication

1 Introduction

Forthcoming human-support robots are anticipated to assist people in a variety of contexts [3, 22] involving socio-emotional personal information (e.g. helping an elderly person live safely independently) best communicated to humans via their innate communication modalities (e.g. speech, facial expressions, gestures), and ideally with established rapport between interlocutors. We aim to continue investigating the introduction of ECAs on social and service robots as a natural user interface metaphor for Human Robot Interaction (HRI) [10, 17].

Establishing and maintaining rapport between humans is a proven determinant of positive communication outcomes, and is the result of a combination of highly socio-cultural-emotional complex processes, some of which are unconscious: *mutual attentiveness* (e.g., mutual gaze, mutual interest, and focus during interaction), *positivity* (e.g., head nods, smiles, friendliness, and warmth) and unconscious *coordination* (e.g., postural mirroring, synchronized movements, balance, and harmony) [11, 26].

In this article, we focus on one of these processes, *coordination*. Specifically, we examine the coordination/mirroring of (1) head movements and (2) facial emotions [2, 8, 14]. Mirroring has been studied extensively in interactions between robots and humans, as well as in interactions between ECAs and humans. However, there are only a few studies [1, 5] that examine interactions between humans and ECAs that are integrated with robots, and none of them examine the effect of mirroring nonverbal behaviors in such interactions.

Our aim is to answer the question of whether integrating an ECA capable of mirroring its interlocutor's head movements and facial emotions (continuously or intermittently) with a support robot will improve the user's experience with that robot, which is capable of performing useful mobile manipulative tasks in a home environment. We review the latest research on our topic and propose to model rapport for human-robot interaction. We discuss the integration of our speaking, expressive, and realistic ECA with a robotic platform, Toyota's Human-Support Robot (HSR, shown in Fig. 1). For the sake of brevity, we will use the term "ECA-HSR" throughout this paper to refer to the integrated ECA and HSR. Furthermore, we discuss how we enabled the ECA-HSR to subtly track its interlocutor's face, and to mirror their head movements and facial emotions in real time. Lastly, we describe a user study we conducted towards answering our research question.

2 Related Work

In human-robot interaction, research on robots establishing rapport is under way, with some research groups investigating verbal [4, 23, 12] and non-verbal [20, 21, 13] social cues. Previous work has examined how mimicry affects human-ECA interaction and human-robot interaction, but not human interaction with an ECA running on a robot. Hasumoto et al. [13] studied the effects of body movement mimicry in human-robot interaction by designing the Reactive Chameleon, a method of generating robot body movements that subtly mimics human body swaying during interactions. They found that subtle mimicry can positively impact the establishment of rapport, while noticeable mimicry can negatively impact it. However, this method was limited to mimicking movements of the torso and did not consider movement of other parts of the robot such as the head.

In the experiment by Riek et al. [20], participants interacted with a robotic chimpanzee named Virgil that exhibited three different types of behavior: full mimicry of head gestures, partial mimicry of nodding gestures only, and no mimicry accompanied by periodic blinks. Afterwards, participants filled out a survey that measured the social attraction toward and emotional credibility of conversation partners. No significant differences were found between participant ratings of the different mimic conditions. However, this might have been due to technical issues, as a few participants "said that the head movements were too erratic or jerky." Other participants wished for the robot to make "non-speech sounds" (backchannel cues) to indicate understanding in conjunction with head gestures. This suggests that robot mimicry of nonverbal behavior by itself might not be enough to create rapport during an interaction with a human.

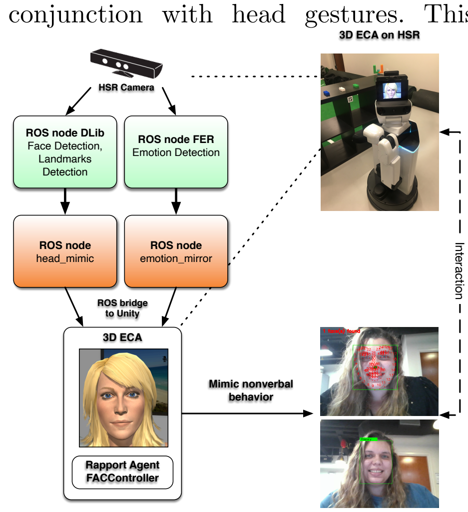

Fig. 1: Data flow in our rapport-building ROS nodes.

Niewiadomski et al. [16] studied how mimicry of smiles influenced interactions between ECAs and humans by testing three different types of ECA smiling behavior when providing backchannel cues: mimicking the smiles of a participant (MS), randomly smiling (RS), and no smiling (NS). They found that participants felt less engaged and more frustrated in condition NS than in condition MS, and "felt more at ease and more listened to" while telling a story to the ECA in condition MS than RS. These results suggest that mimicry in the smiling behavior of an ECA influences "the quality, ease, and warmth, of the user-agent interaction."

In the first experiment of the case study by Stevens et al. [24], participants read some sentences and then listened to an ECA speak some of those sentences incorrectly. They were then asked to say the correct version of the sentence to the ECA. Afterwards, when interacting with the experimental group, the ECA repeated what the subject said while mirroring the eyebrow raises and head nods that were observed during the subject's reading and correcting of potentially erroneous sentences, whereas in the control group no mimicry occurred when the sentence was repeated. The results showed that more prominent cues lead to higher ratings of ECA lifelikeness in the mimic condition, which supports the use of mimicry for building rapport in human-ECA interaction.

Although there exists previous work that integrates an avatar on a robot, to the best of our knowledge no study examines how the mimicry of nonverbal behavior by an ECA running on a robot influences interactions with humans. For example, Domingo et al. [5] projected an avatar on a robotic head and designed a gaze control system that enabled the robotic head to reorient its position based on the location of people that it interacted with. However, the study did not investigate mimicry of nonverbal behavior by the avatar or by the robot.

3 Technical Approach

We integrated the HSR [28] with a fully autonomous ECA. HSR is a social robot designed to assist people with disabilities and the elderly with household tasks such as cleaning or bringing objects. The software empowering the behavior of the HSR runs on top of the Robot Operating System (ROS), a middle-ware for robotic platforms where running programs called *nodes* communicate with each other by subscribing and publishing to data channels called *topics*. The hardware of the HSR includes a wide array of sensors that provide rich data on its surrounding environment, which can be accessed through pre-defined ROS topics. For example, the image data from the robot head camera can be accessed via a ROS topic called */hsrb/head_rgbd_sensor/rgb/image_raw/compressed*.

Our system consists of four ROS nodes and an ECA from a modular framework called eEVA [19], which enables the creation of ECA dialogs suitable for a wide range of scenarios. Two of the nodes run Dlib and the Facial Emotion Recognition (FER) toolkit [25], which are third-party libraries that we adapted for the purpose of analyzing images of faces. The other two nodes, named *head_mimic* and *emotion_mirror*, were created by our group to drive the nonverbal behavior of the ECA-HSR based on the analysis of face images. Below we describe in detail the functionality of each of these ROS nodes.

3.1 Face Detection and Posture Mimicking

To detect the face of a participant interacting with the HSR, we use an adapted version of Dlib [15] that runs on a ROS node and gets images from HSR's Asus Xtion Pro RGB-D camera at approx. 30fps. When the node running Dlib detects a face in the image, it draws a box around the face and marks the face with 68 landmarks as seen in Fig. 1. Then, the node gets the central position of the user's face and publishes that information to a ROS topic. Our *head_mimic* node subscribes to this topic and uses the information to drive the posture mimicry behavior of the ECA-HSR.

To move the HSR's head such that it mirrors the user's head movements, we calculate the delta values of the (x, y) coordinates of the center of the user's face from image to image. Then we scale the delta values (under the assumption that the user is approximately 60 cm away from the robot) and use the scaled values to move the HSR's head, which can move left-right-up-down (all 4 directions and their mixtures). Specifically, it can tilt up or tilt down up to ± 23°, and rotate left or rotate right up to ± 35°. This allows for following the position of the user in the space and the head movements such as "nodding" and "shaking". To move the ECA, which runs on as a stand-alone Unity application, we use the Rosbridge 2[3] library to facilitate communication between the *head_mimic* node and the ECA, and utilize the scaled delta values to move the ECA in the same way as was done for the HSR's head.

[3] http://wiki.ros.org/rosbridge_suite

3.2 Facial Emotion Mirroring

In addition to mimicking posture, the ability to mirror the emotions from facial expressions, or at least the most universal of such emotions, helps to build rapport in human-human communication. According to Ekman [7], there are seven basic emotions that can be expressed by the human face. Our eEVA agent [19] (cf. 4) is an ECA capable of portraying any of these emotions in real time through movements of all the individual facial action units described in Ekman's Facial Action Coding System (FACs) [6].

To enable our system to recognize emotions, we adapted the FER toolkit to run on a ROS node. FER leverages deep learning with Linear Support Vector Machines to classify Ekman's seven basic emotions from facial expressions (see Fig. 1). We also considered another toolkit for emotion classification called EmoPy [18], but ultimately we chose FER because it classified emotions more accurately than EmoPy did. The node running FER publishes the classified emotion to another ROS topic, which is then subscribed to and used by our *emotion_mirror* node to drive the ECA's facial emotion mirroring (which we will refer to as "emotion mirroring" for short) behavior during its interaction with a participant.

3.3 Technical Considerations

Prior to our study, we tested our system and found that the ECA-HSR was performing well under three conditions: (1) the face of a participant was unobstructed, (2) the participant was inside of the maximal scope of sight of the robot's camera, and (3) the participant was facing towards the robot. Dlib excels at detecting faces in a frontal profile but when a participant turns away, it cannot properly detect their face. The same conditions apply when using FER for detecting emotions on facial expressions. Furthermore, we discovered some latency issues when the posture mimicking behavior was enabled, which was likely due to latency issues in the WiFi connection.

4 Experiments and Discussion

Overview. We designed three within-subjects experiments, which are described in the following sections, to assess the impact of various nonverbal behavior skills on the user's sense of comfort and naturalness during their interaction with the robot: *experiment 1* assesses the impact of posture mimicking; *experiment 2* assesses the impact of emotion mirroring; and *experiment 3* assesses the impact of combining posture mimicking with emotion mirroring. With a prior Institutional Review Board (IRB) approval for the study (no. 20210324), we conducted these experiments and discussed the results below.

Material. All three experiments were performed under the following setup: FER and Dlib run on an HP Spectre laptop with 16 GB RAM and four CPUs (Intel Core i7-6500U @ 2.50GHz). The ECA was created using Unity and runs on HSR, which has an Intel Core i7-4700EQ and a NVIDIA Jetson GPU. Communication between the laptop and HSR was facilitated using a 5G WiFi network with an average package return time of 5 ms.

Participants. All three experiments were conducted with 32 participants recruited from the university campus with ages ranging from 17 to 67 years. Participants were 57.6% male and 42.4% female. The participants reported their age-ranges as under 18 (3.0%), 18–24 (45.5%), 25–50 (33.3%), 51–64 (12.1%), and 65 or above (6.1%). They reported their race as White (49%), Asian (24%), African Descent (15%), Indo-Caribbean (3%), Middle-Eastern (3%), and preferred not to identify (6%). The participants reported their ethnicity as Not Hispanic/Latino (69.7%), Hispanic/Latino (27.3%), and 3% preferred not to answer. Their education level was High School (3%), some College (27.3%), Bachelor's Degree (30.3%), some Graduate School (9.1%), Master's Degree (21.2%), Doctoral Degree (6.1%), and Not Specified (3%).

Participants also answered questions, on a 5-point Likert scale, about their previous technical experience: 6.1% claimed little experience, 36.4% had some experience, and 57.6% had a lot or a professional level experience with computers. Then, when asked about their experience with digital avatars and robots, the majority (over 70%) of participants claimed to have little or no previous experience. Thus, while most of the participants had at least some experience with computers, most had little to no experience with virtual agents or robots.

Procedure. Participants were recruited on campus by staff members who verbally asked if they would like to participate in a study where they would interact with a robot. Prior to the experiments, participants were asked to complete consent forms, demographics questionnaire and the pre-experiment questionnaire relating to their previous experience with computers, robots and virtual agents. They were also briefed about the conversational and nonverbal behavior skills of the ECA-HSR. All three experiments were conducted with the ECA running on the HSR as shown in Fig. 1. The HSR was set up in one spot and lighting conditions remained unchanged. For each experiment, participants were asked to stand in front of the robot (at a distance of approximately 60 cm) and were told they could move around within a marked area while the robot's head camera tracked their face. Furthermore, an additional ROS node controlling the speech component of the interaction was enabled, and initiated a conversation with the participant regarding the weather.

Each experiment was introduced by the ECA to differentiate where each condition of nonverbal interaction was present. For example, at the start of first condition of experiment 1, the ECA would announce, "I will first follow your head movements with the character on the screen." As the condition of the nonverbal interaction was announced by the ECA, a staff member would manually activate the appropriate ROS nodes for that mode of interaction. For each condition, participants were allowed to take two turns to ask the ECA-HSR questions related to the weather.

For each experiment, after interacting with the ECA-HSR, participants filled a questionnaire – their choice of either a form on paper or on a computer – to evaluate the ECA-HSR's performance, all without staff supervision. After completing all three experiments, participants completed an additional questionnaire about their experience and the ECA-HSR's overall performance. They were also given the option to provide suggestions for tasks they would like the robot to do in the future and for changes in the physical appearance of the robot.

EXPERIMENT 1: Assessing skills in posture mimicking. We tested the effect on the user of the posture mimicking skills of the ECA alone, of the robot head alone, and of both the ECA and robot head in synchrony. The ROS node running Dlib was used to control the posture mimicking behavior, specifically mirroring of user head movements (Fig. 1). Our independent variable was posture mimicking, with three possible conditions:

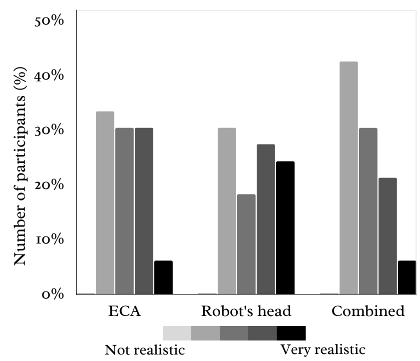

Fig. 2: How realistic were head movements.

1. the ECA mirrors the user's head movements, while the robot head stays immobile;
2. the robot head mirrors the user's head movements, while the ECA stays immobile in the center of the robot screen;
3. both the ECA and the robot head mirror the user's head movements.

Results. A majority of participants (63.6%) identified the interaction with the ECA-HSR to be the most natural in the case where both the robot head and the ECA were mirroring their head movements; furthermore, 38% of those participants commented that they felt as if the ECA-HSR agent was listening to them and that the combined movement gave it more "dimension".

Participants were then asked follow-up questions regarding how realistic the movements of each condition were. The movements were rated on a 5-point Likert scale (1 – not realistic to 5 – very realistic). The realism of the movements by the ECA alone was rated 4 or 5 points by 36.4% participants, at 3 by 30.3%, and at 2 by 33.3% participants. The realism of the movements by the robot's head alone was rated 4 or 5 points by 51.5% of participants, at 3 by 18.2% of participants and at 2 by 30.3% of participants. The realism of the combined movements of the ECA and the robot head was rated 4 or 5 points by 27.3% participants, at 3 by 30.3% of participants and at 2 by 42.4% of participants. None of the movements were rated at 1 – not realistic. Overall, as shown in Fig.2 the participants rated the movements of the robot's head as being more realistic than that of the ECA alone, or that of the ECA and robot head combined.

EXPERIMENT 2: Assessing skills in mirroring facial expressions of emotion. In this experiment, only the performance of the ECA's emotion mirroring skills was evaluated. The ROS node running FER was used to control the mirroring of facial expressions of emotion (Fig. 1, bottom-right). Our independent variable was emotion mirroring, with two possible conditions:

1. mirroring of facial expressions of emotion enabled on ECA;
2. mirroring of facial expressions of emotion displayed on ECA.

Results. After the interaction, the participants were asked which mode of the interaction they found more engaging. 51.5% of participants said the ECA's mirroring of facial expressions made the robot more engaging, while 48.5% said the mirroring did not make the robot more engaging. Participants were then asked two follow-up questions that were rated on a 5-point scale (1 – not realistic/accurate to 5 – very realistic/accurate). The realism of the facial expressions (cf. fig. 4) of the ECA was rated at 4 or 5 points by 21.3% participants, at 3 by 27.3%, and at 2 or 1 by 51.5%, and at 1 by 24.2% participants. The realism of the movements by the mirrored emotions portrayed by the ECA was rated 4 or 5 points by 21.2% of participants, at 3 by 18.2% of participants, at 2 or 1 by 60.6% of participants, as shown in Fig.3. When given an option to explain their choice, participants who favored the ECA's emotion mirroring behavior wrote that it was *nice* or *amusing* although sometimes inaccurate. Those who did not favor described it as *grotesque* and inaccurate, but suggested they would be more receptive to a version that mirrored facial expressions precisely. The responses to those questions confirmed earlier comments regarding the realism and accuracy of portrayed emotions. For both questions, about half of the participants gave a rating of 1 or 2 points. This lead us to conclude that a better performing emotion mirroring module would be beneficial, whereas if this issue is not addressed, it would hinder our efforts to introduce an emotive ECA on a robotic platform.

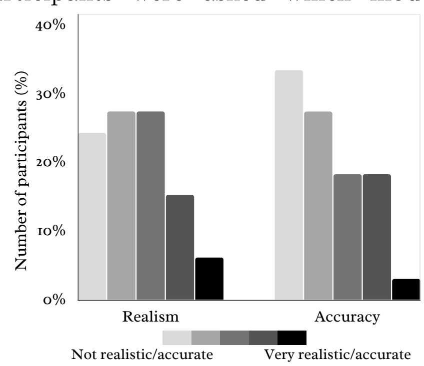

Fig. 3: Realism and accuracy of facial emotion mirroring.

EXPERIMENT 3: Assessing skills combining emotion mirroring with posture mimicking. In this experiment, we tested whether posture mimicking *in conjunction with* emotion mirroring, improves the user's comfort level during human-robot interaction. Our independent variable was combined posture mimicking and emotion mirroring. Both posture mimicking and emotion mirroring were engaged – the ECA mirrored facial emotions, while both the ECA and the robot's head mirrored the movements of the participant's face – for a single turn.

Results. A majority of participants (66.7%) preferred the condition where both posture mimicking and emotive mirroring were enabled. They explained in their questionnaire answers that the two behaviors made the ECA-HSR more engaging, but wished for the emotion mirroring to be more accurate and suffer less latency. Among the remaining participants (33.3%), 73% said they did not prefer this condition because the emotion mirroring was not accurate, but would like it if it was more accurate. Finally, there were two participants (6.2%) who deemed the avatar with its emotion mirroring to be scary and would prefer to not interact with it at all.

Discussion. Before interaction, participants were asked four questions to understand their attitude towards interacting with virtual agents and robots. The questions were:

1. How likely are you to seek interaction with robots or virtual agents?
2. How would you scale your attitude towards interaction with virtual agents?
3. How would you scale your attitude towards interaction with robots?
4. How frustrating do you find interaction with virtual agents or robots?

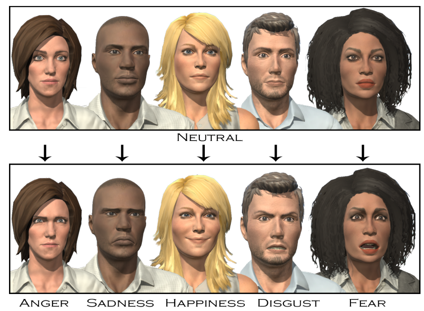

Fig. 4: Examples of eEVA facial expressions

For the first three questions, on a scale from 1 (never/I dislike it) to 5 (very likely/I love it), the percentage of participants that rated it 4 or 5 points was 42.5%, 39.4%, 36.4% respectively, while the percentage that rated it 3 points (neutral stance) was 27.3%, 39.4%, and 39.4% respectively.

For the fourth question, on a scale from 1 (not frustrating) to 5 (very frustrating), 33.4% rated it 1 or 2 points and 42.4% rated it 3 points. In each of these questions, at least 70% of the participants had a neutral or positive attitude towards interacting with robots or virtual agents.

Based on the evaluation of the participants' answers, the inclusion of the ECA on the HSR was favored. As previously mentioned, the majority of participants chose the interaction with both posture mimicking and emotion mirroring. When asked, on a scale of 1 to 5, of their preference to have the ECA on the robot or not, 72.7% of participants chose the two highest scores (4 and 5) meaning *I would prefer a robot with an ECA character*, 6.1% chose a neutral option (3), and 21.2% leaned towards a robot without an ECA (scores 1 and 2).

When participants were asked (on a 5-point Likert scale where 1 meant *very hard* and 5 meant *very easy*) if the interaction with the ECA-HSR was easy, 54.5% rated it with 4 or 5 points while 15.1% rated it with 1 or 2 points. A majority of the participants liked the robot, with 54.5% rating it 4 or 5 points, on a 5-point scale where 1 meant *not at all* and 5 meant *very much*. Overall, we noticed the following patterns: participants with previous experience with robots and/or virtual agents were more critical and had higher expectations of the overall ECA-HSR performance. Furthermore, these participants almost always preferred the integrated of the ECA with the robot.

Another phenomena we investigated was whether participants, given the choice, prefer to interact with an ECA of the same gender as themselves, in the context of a social interaction at home with the ECA-HSR. Table 1 details

Table 1: Participant and selected ECA's gender comparison

Participant's Gender	Selected Gender Before Seeing Images of ECAs			Selected Gender After Seeing Images of ECAs		
	Same	Different	No preference	Same	Different	No preference
Female	9 (64.3%)	5 (35.7%)	0	8 (57.1%)	6 (42.9%)	0
Male	4 (22.2%)	7 (38.9%)	7 (38.9%)	3 (16.7%)	13 (72.2%)	2 (11.1%)

participants' answers to the preference of an ECA's gender, from both selecting gender before seeing images of (25 different) ECAs and from selecting gender after seeing images of ECAs. In regards to the selected gender (of the ECA) before seeing images of ECAs, we found that 9 out of the 14 female participants (64.3%) expressed a preference to interact with an ECA of the same gender while only 4 out of the 18 male participants (22.2%) expressed a preference to interact with an ECA of the same gender. In regards to the selected gender (of the ECA) after seeing images of ECAs, we found that 8 out of the 14 female participants (57.1%) expressed a preference to interact with an ECA of the same gender while only 3 out of the 18 male participants (16.7%) expressed a preference to interact with an ECA of the same gender. Overall, the participants had a slight preference for interacting with a female ECA, as 16 out of the 32 participants (50%) expressed a preference to interact with a female ECA before seeing images of ECAs and 19 out of the 32 participants (59.4%) expressed a preference to interact with a female ECA after seeing images of ECAs. As an additional finding for further investigation, when asked an open-ended questions for suggestions on "any other ideas for robot tasks", all participants, regardless of their prior experience with robots and/or virtual agents, mentioned they would like the robot to clean for them.

The video demonstration of selected fragments of the study experiments can be found at https://tinyurl.com/pyrrmxft.

5 Conclusion

We studied the effects of integrating an ECA capable of mirroring the interlocutor's head movements and facial emotions (continuously or intermittently) with a human-service robot. We found that the integration of the ECA and its nonverbal behaviors has the potential to improve the overall interaction between human and robot. Specifically, users preferred the integrated ECA-HSR over the HSR alone, even though they rated the mirroring of their head movements by the HSR's head alone as being more accurate than that of the mirroring by both the ECA and the HSR. However, occasional inaccurate portrayals of nonverbal behavior, as exemplified by the facial emotion mirroring behavior of the ECA in our study, can contradict user expectations of how the system should behave and thus leave a negative impression of the interaction. One potential explanation for the ECA's inaccurate facial emotion mirroring behavior is that the seven emotions the ECA mirrored were too broad of categories that prevented the mirroring of more nuanced facial expressions. For future work, we plan to have the ECA mirror the individual facial action units of a user's facial expressions and see if this improves the users' perception of the accuracy of the ECA's facial emotion mirroring skills. We also plan to improve upon our methods for measuring the rapport-building capabilities of the ECA-HSR by incorporating validated questionnaires for HRI [27] and ECAs [9]. Finally, we acknowledge the need to improve our experimental methodology for future studies by randomizing stimuli to account for a potential learning effect among participants.

One novel finding is that the participants, regardless of their gender, responded with a slight preference for interacting with a female ECA in the context of a social interaction in a home environment. Another interesting finding is that when asked for "any other ideas about robot tasks", all the participants mentioned they would like the robot to clean for them. While they are not the main focus of this study on rapport-building, these findings point to a future research direction to investigate user preferences regarding the ECA-HSR. For example, future work based on these findings could investigate if people feel more comfortable interacting with the ECA-HSR in a home environment when the ECA is of a specific gender, as well as how much autonomy people feel comfortable giving to a service robot to perform a cleaning task in their homes.

References

1. Cavedon, L., Kroos, C., Herath, D., Burnham, D., Bishop, L., Leung, Y., Stevens, C.J.: "C'Mon dude!": Users adapt their behaviour to a robotic agent with an attention model. International Journal of Human-Computer Studies **80**, 14–23 (2015)
2. Chartrand, T.L., Maddux, W.W., Lakin, J.L.: Beyond the Perception-Behavior link: The Ubiquitous Utility and Motivational Moderators oi Nonconscious Mimicry. The new unconscious pp. 334–361 (2005)
3. Dahl, T.S., Boulos, M.N.K.: Robots in health and social care: A complementary technology to home care and telehealthcare? Robotics **3**(1), 1–21 (2014)
4. Dieter, J., Wang, T., Chaganty, A.T., Angeli, G., Chang, A.: Mimic and Rephrase: Reflective listening in open-ended dialogue. In: Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL). pp. 393–403 (2019)
5. Duque-Domingo, J., Gómez-García-Bermejo, J., Zalama, E.: Gaze Control of a Robotic Head for Realistic Interaction With Humans. Frontiers in Neurorobotics **14**, 34 (2020)
6. Ekman, P.: Facial Action Coding System (FACS). In: A Human Face (2002), nII Article ID (NAID): 10025007347
7. Ekman, P.: An argument for basic emotions. Cognition & Emotion **6**(3/4), 169–200 (1992)
8. Fischer, A., Hess, U.: Mimicking Emotions. Current Opinion in Psychology **17**, 151–155 (2017)
9. Fitrianie, S., Bruijnes, M., Bönsch, A., Brinkman, W.P.: The 19 unifying questionnaire constructs of artificial social agents: An IVA community analysis. pp. 1–8 (10 2020). https://doi.org/10.1145/3383652.3423873
10. Goodrich, M.A., Schultz, A.C.: Human–Robot Interaction: A Survey. Human–Computer Interaction **1**(3), 203–275 (2007)

11. Grahe, J.: The Importance of Nonverbal Cues in Judging Rapport. Journal of Nonverbal behavior **23**(4), 253–269 (1999)
12. Grigore, E.C., Pereira, A., Zhou, I., Wang, D., Scassellati, B.: Talk to Me: Verbal Communication Improves Perceptions of Friendship and Social Presence in Human-Robot Interaction. In: International conference on intelligent virtual agents. pp. 51–63. Springer (2016)
13. Hasumoto, R., Nakadai, K., Imai, M.: Reactive Chameleon: A Method to Mimic Conversation Partner's Body Sway for a Robot. International Journal of Social Robotics **12**(1), 239–258 (2020)
14. Hess, U., Fischer, A.: Emotional mimicry: Why and when we mimic emotions. Social and personality psychology compass **8**(2), 45–57 (2014)
15. King, D.E.: Dlib-ml: A Machine Learning Toolkit. The Journal of Machine Learning Research **10**, 1755–1758 (2009)
16. Niewiadomski, R., Prepin, K., Bevacqua, E., Ochs, M., Pelachaud, C.: Towards a Smiling ECA: Studies on Mimicry, Timing and Types of Smiles. In: Proceedings of the 2nd International Workshop on Social Signal Processing. p. 65–70. SSPW '10, Association for Computing Machinery, New York, NY, USA (2010)
17. Peña, P., Polceanu, M., Lisetti, C., Visser, U.: eEVA as a Real-Time Multimodal Agent Human-Robot Interface. In: Robot World Cup. pp. 262–274. Springer (2018)
18. Perez, A.: EmoPy: A Machine Learning Toolkit for Emotional Expression, 2018, https://tinyurl.com/y6uwnet3
19. Polceanu, M., Lisetti, C.: Time to Go ONLINE! A Modular Framework for Building Internet-Based Socially Interactive Agents. In: Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents. p. 227–229. IVA '19, Association for Computing Machinery, New York, NY, USA (2019)
20. Riek, L.D., Paul, P.C., Robinson, P.: When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. Journal on Multimodal User Interfaces **3**(1), 99–108 (2010)
21. Ritschel, H., Aslan, I., Mertes, S., Seiderer, A., André, E.: Personalized synthesis of intentional and emotional non-verbal sounds for social robots. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). pp. 1–7. IEEE (2019)
22. Šabanović, S.: Robots in society, society in robots. International Journal of Social Robotics **2**(4), 439–450 (2010)
23. Seo, S.H., Griffin, K., Young, J.E., Bunt, A., Prentice, S., Loureiro-Rodríguez, V.: Investigating people's rapport building and hindering behaviors when working with a collaborative robot. IJSR **10**(1), 147–161 (2018)
24. Stevens, C.J., Pinchbeck, B., Lewis, T., Luerssen, M., Pfitzner, D., Powers, D.M., Abrahamyan, A., Leung, Y., Gibert, G.: Mimicry and expressiveness of an ECA in human-agent interaction: familiarity breeds content! Computational cognitive science **2**(1), 1–14 (2016)
25. Tang, Y.: Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239 (2013)
26. Tickle-Degnen, L., Rosenthal, R.: The nature of rapport and its nonverbal correlates. Psychological Inquiry **1**(4), 285–293 (1990)
27. Yagoda, R., Gillan, D.: You want me to trust a ROBOT? The development of a human-robot interaction trust scale. IJSR **4** (08 2012)
28. Yamamoto, T., Terada, K., Ochiai, A., Saito, F., Asahara, Y., Murase, K.: Development of Human Support Robot as the research platform of a domestic mobile manipulator. ROBOMECH Journal **6**(1), 4 (Apr 2019)